

SEQUENCE LISTING

SEQ ID NO: 1
HDEL (peptide)

5 SEQ ID NO: 2
Pichia pastoris OCH1 gene

1	AGATCTGCCT	GACAGCCTTA	AAGAGCCCGC	TAAAAGACCC	GGAAAACCGA	GAGAACTCTG
61	GATTAGCAGT	CTGAAAAAGA	ATCTTCACTC	TGTCTAGTGG	AGCAATTAAT	GTCTTAGCGG
121	CACTTCCTGC	TACTCCGCCA	GCTACTCCTG	AATAGATCAC	ATACTGCAAA	GACTGCTTGT
10 181	CGATGACCTT	GGGGTTATTT	AGCTTCAAGG	GCAATTTTTG	GGACATTTTG	GACACAGGAG
241	ACTCAGAAAC	AGACACAGAG	CGTTCTGAGT	CCTGGTGCTC	CTGACGTAGG	CCTAGAACAG
301	GAATTATTGG	CTTTATTTGT	TTGTCCATTT	CATAGGCTTG	GGTAATAGA	TAGATGACAG
361	AGAAATAGAG	AAGACCTAAT	ATTTTTTGTT	CATGGCAAAT	CGCGGGTTCG	CGGTCGGGTC
421	ACACACGGAG	AAGTAATGAG	AAGAGCTGGT	AATCTGGGGT	AAAAGGGTTC	AAAAGAAGGT
15 481	CGCCTGGTAG	GGATGCAATA	CAAGGTTGTC	TTGGAGTTTA	CATTGACCAG	ATGATTTGGC
541	TTTTTCTCTG	TTCAATTCAC	ATTTTTCAGC	GAGAATCGGA	TTGACGGAGA	AATGGCGGGG
601	TGTGGGGTGG	ATAGATGGCA	GAAATGCTCG	CAATCACCGC	GAAAGAAAGA	CTTTATGGAA
661	TAGAACTACT	GGGTGGTGTA	AGGATTACAT	AGCTAGTCCA	ATGGAGTCCG	TTGGAAAGGT
721	AAGAAGAAGC	TAAAACCGGC	TAAGTAACTA	GGGAAGAATG	ATCAGACTTT	GATTTGATGA
20 781	GGTCTGAAAA	TACTCTGCTG	CTTTTTCAGT	TGCTTTTTC	CTGCAACCTA	TCATTTTCCT
841	TTTCATAAGC	CTGCCTTTTC	TGTTTTCAGT	TATATGAGTT	CCGCCGAGAC	TTCCCCAAAT
901	TCTCTCCTGG	AACATTCTCT	ATCGCTCTCC	TTCCAAGTTG	CGCCCCCTGG	CAGTCCCTAG
961	TAATATTACC	ACGCGACTTA	TATTCAGTTC	CACAATTTCC	AGTGTTCTGA	GCAAATATCA
1021	TCAGCCATGG	CGAAGGCAGA	TGGCAGTTTG	CTCTACTATA	ATCCTCACAA	TCCACCCAGA
25 1081	AGGTATTACT	TCTACATGGC	TATATTCGCC	GTTTCTGTCA	TTTGCCTTTT	GTACGGACCC
1141	TCACAACAAT	TATCATCTCC	AAAAATAGAC	TATGATCCAT	TGACGCTCCG	ATCACTTGAT
1201	TTGAAGACTT	TGGAAGCTCC	TTACAGTTG	AGTCCAGGCA	CCGTAGAAGA	TAATCTTCGA
1261	AGACAATTGG	AGTTTCATTT	TCCTTACCGC	AGTTACGAAC	CTTTTCCCCA	ACATATTTGG
1321	CAAACGTGGA	AAGTTTCTCC	CTCTGATAGT	TCCTTTCCGA	AAAACCTCAA	AGACTTAGGT
30 1381	GAAAGTTGGC	TGCAAAGGTC	CCCAAATTAT	GATCATTTTG	TGATAACCGA	TGATGCAGCA
1441	TGGGAACTTA	TTACCATGA	ATACGAACGT	GTACCAGAAG	TCTTGGAAGC	TTTCCACCTG
1501	CTACCAGAGC	CCATTCTAAA	GGCCGATTTT	TTCAGGTATT	TGATTCTTTT	TGCCCCGTGA
1561	GGACTGTATG	CTGACATGGA	CACTATGTGA	TTAAAACCAA	TAGAATCGTG	GCTGACTTTC
1621	AATGAACTA	TTGGTGGAGT	AAAAACAAT	GCTGGGTTGG	TCATTGGTAT	TGAGGCTGAT
35 1681	CCTGATAGAC	CTGATTGGCA	CGACTGGTAT	GCTAGAAGGA	TACAATTTTG	CCAATGGGCA
1741	ATTCAGTCCA	AACGAGGACA	CCCAGCACTG	CGTGAAGTGA	TTGTAAGAGT	TGTCAGCACG
1801	ACTTTACGGA	AAGAGAAAAG	CGGTTACTTG	AACATGGTGG	AAGGAAAGGA	TCGTGGAAGT
1861	GATGTGATGG	ACTGGACGGG	TCCAGGAATA	TTTACAGACA	CTCTATTTGA	TTATATGACT
1921	AATGTCAATA	CAACAGGCCA	CTCAGGCCAA	GGAAATTGGAG	CTGGCTCAGC	GTATTACAAT
40 1981	GCCTTATCGT	TGGAAGAACG	TGATGCCCTC	TCTGCCCGCC	CGAACGGAGA	GATGTTAAAA
2041	GAGAAAAGTCC	CAGGTAAATA	TGCACAGCAG	GTTGTTTTAT	GGGAACAATT	TACCAACCTG
2101	CGCTCCCCCA	AATTAATCGA	CGATATTCTT	ATTCTTCCGA	TCACCAGCTT	CAGTCCAGGG
2161	ATTGGCCACA	GTGGAGCTGG	AGATTGTGAAC	CATCACCTTG	CATATATTAG	GCATACATTT
2221	GAAGGAAGTT	GGAAGGACTA	AAGAAAGCTA	GAGTAAATA	GATATAGCGA	GATTAGAGAA
45 2281	TGAATACCTT	CTTCTAAGCG	ATCGTCCGTC	ATCATAGAAT	ATCATGGACT	GTATAGTTTT
2341	TTTTTTGTAC	ATATAATGAT	TAAACGGTCA	TCCAACATCT	CGTTGACAGA	TCTCTCAGTA
2401	CGCGAAATCC	CTGACTATCA	AAGCAAGAAC	CGATGAAGAA	AAAAACAACA	GTAACCCAAA
2461	CACCACAACA	AACACTTTAT	CTTCTCCCCC	CCAACACCAA	TCATCAAAGA	GATGTCGGAA
2521	CACAAACACC	AAGAAGCAAA	AACTAACCCC	ATATAAAAAC	ATCCTGGTAG	ATAATGCTGG
50 2581	TAACCCGCTC	TCCTTCCATA	TTCTGGGCTA	CTTCACGAAG	TCTGACCGGT	CTCAGTTGAT
2641	CAACATGATC	CTCGAAATGG	GTGGCAAGCA	TCGTTCCAGA	CCTGCCTCCT	CTGGTAGATG
2701	GAGTGTGTGT	TTTGACAGGG	GATTACAAGT	CTATTGATGA	AGATACCCTA	AAGCAACTGG
2761	GGGACGTTCC	AATATACAGA	GACTCCTTCA	TCTACCAGTG	TTTGTGTGCAC	AAGACATCTC
55 2821	TTCCCATTTGA	CACTTTCCGA	ATTGACAAGA	ACGTCGCAC		

SEQ ID NO: 3
Pichia pastoris alpha-1,6-mannosyltransferase encoded by the OCH1 gene
(nt 1027-2241)

MAKADGSLLYNPHNPPRRYYFYMAIFAVSVICVLYGPSQQLSSPKIDYDPLTLRSLDLKTLLEAPSQLSPG
TVEDNLRRQLEFHFYPYRSYEPFPQHIWQTKVSPSDSSFKNFKDLGESWLQSPNYDHFVI PDAAWELI
HHEYERVPEVLEAFHLLPEPILKADFFRYLILFARGGLYADMDTMLLKPIESWLTFNETIGGVKNNAGLVI
GIEADPDRPDWHDWYARRIQFCQWAIQSKRGHPALRELIVRVSTTLRKEKSGYLMNVEGKDRGSDVMDWT
5 GPGIFTDTLFDYMTNVNTTGHSGQGIGAGSAYYNALSLEERDALSARPNGEMLKEKVPGKYAQQVVLWEQF
TNLRSPKLIDDILILPITSFSPGIGHSGAGDLNHHLAYIRHTFEGSWKD

SEQ ID NO: 4

5' GGAATTCAGCATGGAGTATGGATCATGGAGTCCGTTGGAAAGG

SEQ ID NO: 5

5' GCCGCTCGAGCTAGCTTTCTTTAGTCC

SEQ ID NO: 6

Plasmid pGlycoSwitchM8 (2875 bp)

AGATCTAACATCCATAATCGATCTAAGCTATATTGCGCGTTTCTGTCAATTTGCGTTTTGTACGGACCCCTCA
CAACAATTATCATCTCCAAAAATAGACTATGATCCATTGACGCTCCGATCACTTGATTTGAAGACTTTGGA
AGCTCCTTCACAGTTGAGTCCAGGCACCGTAGAAGATAATCTTGAAGACAATTGGAGTTTCATTTTCCTT
ACCGCAGTTACGAACCTTTTCCCAACATATTTGGCAAACGTGGAAAGTTTCTCCCTCTGATAGTTCTTTT
20 CCGAAAAAATTCAAAGACTTAGGTGAAAGTTGGCTGCAAAGGTCCCAAAATTATGATCATTTTGTGATACC
CGATGATGCAGCATGGGAACCTTATTACCATGAATACGAACGTGTACCAGAAGTCTTGAAGCTTTTGATT
TTAACGACTTTTAAACGACAACCTTGAGAAGATCAAAAAACAATAATTATTGCGGAAACGAGGAATTACGT
GGCCAGCCGCGCTCTCGGATCGGTACCTCGAGCCGCGCGCGCCAGCTTTCTAGAGAACAAAACTC
ATCTCAGAAGAGGATCTGAATAGCGCCGTCGACCATCATCATCATCATCATTGAGTTTGTAGCCTTAGACA
25 TGACTGTTCTCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTCTTGCTAGATTCTAATCAAGAGGATGT
CAGAATGCCATTTGCCTGAGAGATGCAGGCTTCATTTTTGATACTTTTTTATTTGTAACCTATATAGTATA
GGATTTTTTTTTGTCAATTTGTTTCTTCTCGTACGAGCTTGCTCCTGATCAGCCTATCTCGCAGCTGATGAA
TATCTTGTGGTAGGGGTTTGGGAAAATCATTGAGTTTGATGTTTTTCTTGGTATTTCCCACTCCTCTTCA
GAGTACAGAAGATTAAAGTGAGACCTTCGTTTGTGCGGATCCCCACACACCATAGCTTCAAAATGTTTCTA
30 CTCCTTTTTTACTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCGTACCACTTCAAAACACCCAAGCACA
GCATACTAAATTTCCCTCTTCTTCTCTAGGGTGTGTTAATTACCCGTACTAAAGGTTTGGAAAAGAA
AAAAGAGACCGCCTCGTTTCTTTTCTTCTGCGAAAGGCAATAAAAAATTTTATCACGTTTCTTTTCT
TGAAATTTTTTTTTTAGTTTTTTTTCTTTTCTTTTCAAGTGACCTCCATTGATATTTAAGTTAATAAAGGTTCTT
AATTTCTCAAGTTTTCAGTTTTCATTTTTCTTGTCTATTACAACCTTTTTTTTACTTTCTTGTTCATTAGAAAGA
35 AAGCATAGCAATCTAATCTAAGGGGCGGTGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAA
TACGACAAGGTGAGGAACTAAACCATGGCCAAGTTGACCAGTGCCGTTCCGGTGCTCACCGCGCGCAGCT
CGCCGGAGCGGTGAGTTCTGGACCGACCGGCTCGGGTTCTCCGGGACTTCGTGGAGGACGACTTCGCCG
GTGTGGTCCGGGACGACGTGACCTGTTTCATCAGCGCGGTCCAGGACCAGGTGGTGCCGGAACAACCTG
GCCTGGGTGTGGGTGCGCGGCCTGGACGAGCTGTACGCCGAGTGGTCCGAGGTGCTGTCCACGAACCTCCG
40 GGACGCCTCCGGGCCGGCCATGACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAGTTGCGCCTGCGCGACC
CGGCCGGAACCTGCGTGCACTTCGTGGCCGAGGAGCAGGACTGACACGTCCGACGGCGGCCACGGGTCCC
AGGCCTCGGAGATCCGTCCCCCTTTTCTTTTGTGATATCATGTAATTAGTTATGTACGCTTACATTAC
GCCCTCCCCCACATCCGCTCTAACCGAAAAGGAAGGAGTTAGACAACCTGAAGTCTAGGTCCCTATTTAT
TTTTTTATAGTTATGTTAGTATTAAGAACGTTATTTATATTTCAAATTTTTCTTTTTTTTTCTGTACAGACG
45 CGTGACGCATGTAACATTATACTGAAAACCTTGCTTGAGAAGGTTTTTGGGACGCTCGAAGGCTTTAATTT
GCAAGCTGGAGACCAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTG
GCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA
CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCC
TGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCAATGCTCAGCGTGT
50 AGGTATCTCAGTTCGGTGTAGGTGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAAGTGGTGCCT
CCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAG
CAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGCCT
AACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAG
AGTTGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGA

GACGCATGTCATGAGATTATTGGAAACCACCAGAATCGAATATAAAAGGCGAACACCTTTCCCAATTTTGG
 TTTCTCCTGACCCAAAGACTTTAAATTTAATTTATTTGTCCCTATTTCAATCAATTGAACAACCTATTTTCG
 GAAACGATGAGATTTCTTCAATTTTTACTGCTGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCTCCAGT
 CAACACTACAACAGAAGATGAAACGGCACAAATTCGGGCTGAAGCTGTATCGGTTACTCAGATTTAGAAG
 5 GGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATACTACT
 ATTGCCAGCATTTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAGAGGCTGAAGCTGAATTCGCCAC
 AAAACGTGGATCTCCCAACCTACGAGGGCGGCAGCAGTCAAGGCCGCATTCCAGACGTGCTGGAACGCTT
 ACCACCATTTTGCCTTTCCCATGACGACCTCCACCCGGTCAGCAACAGCTTTGATGATGAGAGAAACGGC
 TGGGGCTCGTCGGCAATCGATGGCTTGGACACGGCTATCCTCATGGGGGATGCCGACATTGTGAACACGAT
 10 CCTTCAGTATGTACCGCAGATCAACTTCACCACGACTGCGGTTGCCAACCAAGGATCCTCCGTGTTTCGAGA
 CCAACATTTCGGTACCTCGGTGGCCTGCTTTCTGCCTATGACCTGTTGCGAGGTCCTTTTCAGCTCCTTGGCG
 ACAACACGACCCCTGGTAAACAGCCTTCTGAGGCAGGCTCAAACACTGGCCAACGGCCTCAAGGTTGCGTT
 CACCATCCCAGCGGTGTCGGGACCTACCGTCTTCTTCAACCCTACTGTCCGGAGAGTGGTGATCTA
 GCAACAACGTGCTGAAATTGGAAGCCTGGTGCTCGAGTGGACACGCTTGAGCGACCTGACGGGAAACCCG
 15 CAGTATGCCAGCTTGCAGCAGAAGGGCGAGTCGTATCTCCTGAATCCAAAGGGAAGCCCGGAGGCATGGCC
 TGGCCTGATTGGAACGTTTGTGACGACGAGCAACGGTACCTTTTCAGGATAGCAGCGGCAGCTGGTCCGGCC
 TCATGGACAGCTTCTACGAGTACCTGATCAAGATGTACCTGTACGACCCGGTTGCGTTTGCACACTACAAG
 GATCGCTGGGTCTTGGTGCCGACTCGACCATTTGGGCATCTCGGCTCTCACCCGTGACGCGCAAGGACTT
 GACCTTTTTTGTCTTTCGTACAACGGACAGTCTACGTGCGCAAACTCAGGACATTTGGCCAGTTTGGCGGTG
 20 GCAACTTCATCTTGGGAGGCATTCTCCTGAACGAGCAAAAGTACATTGACTTTGGAATCAAGCTTGCCAGC
 TCGTACTTTGGCACGTACACCCAGACGGCTTCTGGAATCGGCCCCGAAGGCTTCGCGTGGGTGGACAGCGT
 GACGGGCGCCGGCGGCTCGCCGCCCTCGTCCAGTCCGGGTTCTACTCGTCGGCAGGATTCTGGGTGACGG
 CACCGTATTACATCCTGCGGCCGGAGACGCTGGAGAGCTTGTACTACGCATACCGCGTCACGGGCGACTCC
 AAGTGGCAGGACCTGGCGTGGGAAGCGTTGAGTGCCATTGAGGACGCATGCCGCGCCGGCAGCGCGTACTC
 25 GTCCATCAACGACGTGACGCAGGCCAACGGCGGGGGTGCCTCTGACGATATGGAGAGCTTCTGGTTTGCCG
 AGGCGCTCAAGTATGCGTACCTGATCTTTGCGGAGGAGTCGGATGTGCAGGTGCAGGCCACCGGCGGGAAC
 AAATTTGTCTTTAACACGGAGGCGACCCCTTTAGCATCCGTTTCATCATCACGACGGGGCGGCCACCTTGC
 TCACGACGAGTTGTAATCTAGGGCGGCCGCCAGCTTTCTAGAGAAACAAAACCTCATCTCAGAAGAGGATCT
 GAATAGCGCCGTCGACCATCATCATCATCATTGAGTTTGTAGCCTTAGACATGACTGTTCTCAGTTTC
 30 AAGTGGGCGGTCGACGAGACCGGTCCTGCTAGATTCTAATCAAGAGGATGTGAGAATGCCATTTGCCT
 GAGAGATGCAGGCTTCATTTTGTATCTTTTATTTTGTAACTATATAGTATAGGATTTTTTTTGTCACTT
 TTGTTTTCTTCTCGTACGAGCTTGCTCCTGATCAGCCTATCTCGCAGCTGATGAATATCTTGTGGTAGGGT
 TTGGGAAAATCATTGAGTTTGTGTTTTTCTTGGTATTTCCCACTCCTCTTCAGAGTACAGAAGATTAAG
 TGAGACCTTCTGTTTGTGCGGATCCCCACACACCATAGCTTCAAAATGTTTTCTACTCTTTTTTTACTCTTC
 35 CAGATTTTCTCGGACTCCGCGCATCGCCGTACCACTTCAAAACACCCAAGCACAGCATACTAAATTTTCCC
 TCTTTCTTCTCTAGGGTGTGTTAATTACCGTACTAAAGGTTTGGAAAAGAAAAAGAGACCGCCTCGT
 TTCTTTTTCTTCTCGTCAAAAAGGCAATAAAATTTTTATCACGTTTCTTTTTCTTGAAATTTTTTTTTTTA
 GTTTTTTTCTCTTTCAGTGACCTCCATTGATATTTAAGTTAATAAACGGTCTTCAATTTCTCAAGTTTCAG
 TTTTCAATTTTCTTGTCTATTACAATTTTTTTTACTTCTTGTTCATTAGAAAGAAAGCATAGCAATCTAAT
 40 CTAAGGGGCGGTGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGACAAGGTGAGGAA
 CTAACCATGGCCAAGTTGACCAAGTGCCTTCCGGTGCTCACCGCGCGGACGTGCGCGGAGCGGTGAGT
 TCTGGACCGACCGGCTCGGGTCTCCCGGACTTCGTGGAGGACGACTTCGCCGGTGTGGTCCGGGACGAC
 GTGACCTGTTTCATCAGCGCGGTCCAGGACCAGGTGGTGCCGGACAACACCTGGCCTGGGTGTGGGTGCG
 CGGCCTGGACGAGCTGTACGCCGAGTGGTGGGAGGTGCTGTCCACGAACCTTCGGGACGCCTCCGGGCCGG
 45 CCATGACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAGTTCGCCCTGCGCGACCCGGCCGGCAACTGCGTG
 CACTTCGTGGCCGAGGAGCAGGACTGACAGCTCCGACGGCGGCCACGGGTCCAGGCCTCGGAGATCCGT
 CCCCCTTTCTTCTGTCATATCATGTAATTAGTTATGTACGCTTACATTACGCCCCTCCCCCACATCC
 GCTCTAACCGAAAAGGAAGGAGTTAGACAACCTGAAGTCTAGGTCCCTATTTATTTTTTATAGTTATGTT
 AGTATTAAGAAGCTTATTTATATTTCAAATTTTTCTTTTTTTCTGTACAGACGCGTGTACGCAATGTAACA
 50 TTATACTGAAAACCTTGCTTGAGAAGGTTTTGGGACGCTCGAAGGCTTTAATTTGCAAGCTGGAGACCAAC
 ATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCT
 CCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA
 GATACCAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATAC
 CTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCCGT

GTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTTCAGCCCGACCGCTGCGCCTTATCCG
GTAACCTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG
ATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAG
AAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGAT
5 CCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAA
GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAGG
GATTTTGGTCATGAGATC

SEQ ID NO: 10
10 *S. cerevisiae* Kre2 protein
MALFLSKRLLRFTVIAGAVIVLLLLTLNSNSRTQQYIPSSIAAFDFTSGSISPEQQVISEENDAKKLEQSAL
NSEASEDSEAMDEESKALKAAAEKADAPIDTKTMDYITPSFANKAGKPKACYVTLVRNKKELKGLLSSIKY
VENKINKKFPYPWVFLNDEPFTEEFKEAVTKAVSSEVKFGILPKEHWSYPEWINQTKAAEIRADAATKYIY
GGSESYRHMCRYQSGFFWRHELLEEDWYWRVEPDIKLYCDINYDVFKWMQENKVGFTVSIHEYEV TIP
15 TLWQTSMDFIKKNPEYLDENNLM SFLSNDNGKTYNLCHFWSNFEIANLNLWRS PAYREYFD TLDHQGGFFY
ERWGDAPVHSIAAALFLPKDKIHYFSDIGYHPPYDNCPLDKEVYNSNNCECDQGNDFTFQGYSCGKEYYD
AQGLVKPKNWKKFRE

SEQ ID NO: 11
20 *S. cerevisiae* Kre2 Golgi localization signal (first 100 amino acids)
MALFLSKRLLRFTVIAGAVIVLLLLTLNSNSRTQQYIPSSIAAFDFTSGSISPEQQVISEENDAKKLEQSA
LNSEASEDSEAMDEESKALKAAAEKADAP

SEQ ID NO: 12
25 Human GnTI cDNA
ATGCTGAAGAAGCAGTCTGCAGGGCTTGTGCTGTGGGGCGCTATCCTCTTTGTGGCCTGGAATGCCCTGCT
GCTCCTCTTCTTCTGGACGCGCCAGCACCTGGCAGGCCACCCTCAGTCAGCGCTCTCGATGGCGACCCCG
CCAGCCTCACCCGGGAAGTGATTTCGCCTGGCCCAAGACGCCGAGGTGGAGCTGGAGCGGCAGCGTGGGCTG
30 CTGCAGCAGATCGGGGATGCCCTGTGAGCCAGCGGGGAGGGTGCCACCCGCGGCCCTCCCGCCAGCC
GCGTGTGCCTGTGACCCCGCGCGCGGCGGTGATTCCTCATCCTGGTCATCGCCTGTGACCGCAGCACTGTTT
GGCGCTGCCTGGACAAGCTGCTGCATTATCGGCCCTCGGCTGAGCTCTTCCCCATCATCGTTAGCCAGGAC
TGCGGGCAGCAGGAGACCGGCCAGGCCATCGCCTCCTACGGCAGCGCGGTACGCACATCCGGCAGCCCGA
CCTGAGCAGCATTGCGGTGCGCGCGGACACCGCAAGTTCAGGGCTACTACAAGATCGCGCGCCACTACC
35 GCTGGGCGCTGGGCCAGGTCTTCCGGCAGTTTCGCTTCCCCGCGGCCGTGGTGGTGGAGGATGACCTGGAG
GTGGCCCCGACTTCTTCGAGTACTTTCCGGCCACCTATCCGCTGCTGAAGGCCGACCCCTCCCTGTGGTG
CGTCTCGGCCTGGAATGACAACGGCAAGGAGCAGATGGTGGACGCCAGCAGGCCTGAGCTGCTCTACCGCA
CCGACTTTTTTCCCTGGCCTGGGCTGGCTGCTGTTGGCCGAGCTCTGGGCTGAGCTGGAGCCCAAGTGGCCA
AAGGCCTTCTGGGACGACTGGATGCGCGCGCCGAGCAGCGGCAGGGGCGGGCCTGCATACGCCCTGAGAT
CTCAAGAACGATGACCTTTGGCCGCAAGGGTGTGAGCCACGGGCAGTTCTTTGACCAGCACCTCAAGTTTA
40 TCAAGCTGAACCAGCAGTTTGTGCACTTCACCCAGCTGGACCTGTCTTACCTGCAGCGGGAGGCCTATGAC
CGAGATTTCTCGCCCGCGTCTACGGTGCTCCCCAGCTGCAGGTGGAGAAAGTGAGGACCAATGACCGGAA
GGAGCTGGGGGAGGTGCGGGTGCAGTATACGGGCAGGGACAGCTTCAAGGCTTTCCGCAAGGCTCTGGGTG
TCATGGATGACCTTAAGTCGGGGGTTCGAGAGCTGGCTACCGGGGTATTGTACCTTCCAGTTCCGGGGC
CGCCGTGTCCACCTGGCGCCCCACCGACGTGGGAGGGCTATGATCCTAGCTGGAATATGCTGAAGAAGCA
45 GTCTGCAGGGCTTGTGCTGTGGGGCGCTATCCTCTTTGTGGCCTGGAATGCCCTGCTGCTCCTCTTCTTCT
GGACGCGCCAGCACCTGGCAGGCCACCCTCAGTCAGCGCTCTCGATGGCGACCCCGCCAGCCTCACCCGG
GAAGTGATTTCGCCTGGCCCAAGACGCCGAGGTGGAGCTGGAGCGGCAGCGTGGGCTGCTGCAGCAGATCGG
GGATGCCCTGTGAGCCAGCGGGGAGGGTGCCACCCGCGGCCCTCCCGCCAGCCGCGTGTGCCTGTGA
CCCCCGCGCCGGCGGTGATTCCTCATCCTGGTCATCGCCTGTGACCGCAGCACTGTTCCGGCGCTGCCTGGAC
50 AAGCTGCTGCATTATCGGCCCTCGGCTGAGCTCTTCCCCATCATCGTTAGCCAGGACTGCGGGCAGCAGGA
GACGGCCAGGCCATCGCCTCCTACGGCAGCGCGGTACGCACATCCGGCAGCCCGACCTGAGCAGCATTG
CGGTGCCCGCGGACACCGCAAGTTCAGGGCTACTACAAGATCGCGCGCCACTACCGCTGGGCGCTGGGC
CAGGTCTTCCGGCAGTTTCGCTTCCCCGCGGCCGTGGTGGTGGAGGATGACCTGGAGGTGGCCCCGACTT
CTTCGAGTACTTTCCGGCCACCTATCCGCTGCTGAAGGCCGACCCCTCCCTGTGGTGGCTCTCGGCCTGGA

5 ATGACAACGGCAAGGAGCAGATGGTGGACGCCAGCAGGCCTGAGCTGCTCTACCGCACCGACTTTTTCCCT
 GGCCTGGGCTGGCTGCTGTTGGCCGAGCTCTGGGCTGAGCTGGAGCCCAAGTGGCCAAAGGCCTTCTGGGA
 CGACTGGATGCGGCGGCCGAGCAGCGGCAGGGGCGGGCCTGCATACGCCCTGAGATCTCAAGAACGATGA
 CCTTTGGCCGCAAGGGTGTGAGCCACGGGCAGTTCTTTGACCAGCACCTCAAGTTTATCAAGCTGAACCAG
 10 CAGTTTGTGCACTTCAACCAGCTGGACCTGTCTTACCTGCAGCGGGAGGCCTATGACCGAGATTTCTTCGC
 CCGCGTCTACGGTGCTCCCCAGCTGCAGGTGGAGAAAGTGAGGACCAATGACCGGAAGGAGCTGGGGGAGG
 TGCGGGTGCAGTATACGGGCAGGGACAGCTTCAAGGCTTTTCGCCAAGGCTCTGGGTGTCTATGGATGACCTT
 AAGTCGGGGGTTCCGAGAGCTGGCTACCGGGGTATTGTACCTTCCAGTTCGGGGGCCCGGTGTCCACCT
 GCGCCCCCACCGACGTGGGAGGGCTATGATCCTAGCTGGAAT

SEQ ID NO: 13

Human GnTI protein

1 MLKKQSAGLV LWGAILFVAW NALLLLFFWT RPAPGRPPSV SALDGDPAVL TREVIRLAQD
 61 AEVELEERQRL LLQQIGDALSL SQRGRVPTAA PPAQPRVPVT PAPAIVIPILV IACDRSTVRR
 15 121 CLDKLLHYRP SAELFPIIVS QDCGHEETAQ AIASYGSALT HIRQPDLSLI AVPPDHRKFQ
 181 GYYKIARHYR WALGQVFRQF RFPAVVVED DLEVAPDFFE YFRATYPLLK ADPSLWCVSA
 241 WNDNGKEQMV DASRPELLYR TDFPGLGLWL LLAEWLAELE PKWPKAFWDD WMRRPEQRQF
 301 RACIRPEISR TMTFGRKQVS HGQFFDQHLK FIKLNQQFVH FTQLDLSYLQ REAYDRDFLA
 361 RVYGAPQLQV EKVRTNDRKE LGEVRVQYTG RDSFKAFKA LGVMDLKLKSG VPRAGYRGIV
 20 421 TFQFRGRRVH LAPPTWEGY DPSWN

SEQ ID NO: 14

pPIC6AKrecoGnTI

25 GAAATTTTTTTTTTTAGTTTTTTCTCTTTCACTGACCTCCATTGATATTTAAGTTAATAAACGGTCTTCA
 ATTTCTCAAGTTTCAGTTTCATTTTTCTTGTTCTATTACAACTTTTTTTTACTTCTTGTTTATTAGAAAGAA
 AGCATAGCAATCTAATCTAAGGGGCGGTGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAAT
 30 ACGACAAGGTGAGGAACATAAACCATGGCCAAGCCTTTGTCTCAAGAAGAATCCACCCTCATTGAAAGAGCA
 ACGGCTACAATCAACAGCATCCCCATCTCTGAAGACTACAGCGTCGCCAGCGCAGCTCTCTCTAGCGACGG
 CCGCATCTTCACTGGTGTCAATGTATATCATTTTTACTGGGGGACCTTGTGCAGAACTCGTGGTGTCTGGGCA
 CTGCTGCTGCTGCGGCAGCTGGCAACCTGACTTGTATCGTCGCGATCGGAAATGAGAACAGGGGCATCTTG
 AGCCCCGCGGACGGTGCCGACAGGTGCTTCTCGATCTGCATCCTGGGATCAAAGCCATAGTGAAGGACAG
 TGATGGACAGCCGACGGCAGTTGGGATTCTGTAATTGCTGCCCTCTGGTTATGTGTGGGAGGGCTAAGCAC
 35 TTCGTGGCCGAGGAGCAGGACTGACACGTCCGACGGCGGCCACGGGTCCCAGGCCTCGGAGATCCGTCCC
 CCTTTTCTTTGTGATATCATGTAATTAGTTATGTACGCTTACATTACGCCCTCCCCCACATCCGCT
 CTAACCGAAAGGAAGGAGTTAGACAACCTGAAGTCTAGGTCCCTATTTATTTTTTATAGTTATGTTAGT
 ATTAAGAACGTTATTTATATTTCAAATTTTTCTTTTTTTCTGTACAGACGCGTGTACGCATGTAACATTA
 TACTGAAAACGTTGCTTGAGAAGTTTTGGGACGCTCGAAGGCTTTAATTTGCAAGCTGGAGACCAACATG
 40 TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCG
 CCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGAT
 ACCAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCCTGCCGCTTACCGGATACCTG
 TCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCCGTGTA
 GGTGCTTCGCTCCAAGCTGGGCTGTGTGCAGAACCCCCCGTTAGCCCGACCGCTGCGCCTTATCCGGTA
 45 ACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATT
 AGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAG
 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCG
 GCAAAACAAACCCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGA
 TCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAACTCACGTTAAGGGAT
 50 TTTGGTCATGAGATCAGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTTGCCATCCGACATCC
 ACAGGTCCATTCTCACACATAAGTGCCAAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTTGCAAC
 GCAGGACCTCCACTCCTCTTCTCCTCAACACCCACTTTTGCCATCGAAAAACCGCCAGTTATTGGGCTT
 GATTGGAGCTCGCTCATTCCAATTCTTCTATTAGGCTACTAACACCATGACTTTATTAGCCTGTCTATCC
 TGGCCCCCTGGCGAGGTTTCATGTTTGTATTTCGGAATGCAACAAGCTCCGCATTACACCCGAACATCA
 55 CTCCAGATGAGGGCTTTCTGAGTGTGGGGTCAAATAGTTTCATGTTCCCAATGGCCCAAACTGACAGT

5 TTAACGCTGTCTTGGAACCTAATATGACAAAAGCGTGATCTCATCCAAGATGAACTAAGTTTGGTTTCGTT
 GAAATGCTAACGGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCGGCATACCGTTTGTCTTGTGGTATT
 GATTGACGAATGCTCAAAAATAATCTCATTAAATGCTTAGCGCAGTCTCTCTATCGCTTCTGAACCCCGGTG
 CACCTGTGCCGAAACGCAATGGGGAAACACCCGCTTTTTGGATGATTATGCATTGTCTCCACATTGTATG
 10 CTTCCAAGATTCTGGTGGGAATACTGCTGATAGCCTAACGTTTATGATCAAAATTTAACTGTTCTAACCCC
 TACTTGACAGCAATATATAAACAGAAGGAAGCTGCCCTGTCTTAAACCTTTTTTTTTTATCATCATTATTAG
 CTTACTTTTCATAATTGCGACTGGTTCCAATTGACAAGCTTTTGATTTTAACGACTTTTAACGACAACCTTGA
 GAAGATCAAAAAACAATAATTATTGAAACGAGGAATTATCCATATTGCAAGCAGTTCCACTCGAAAGC
 ATGGCCCTCTTTCTCAGTAAGAGACTGTTGAGATTTACCGTCATTGCAGGTGCGGTATTGTTCTCCTCCT
 AACATTGAATCCAACAGTAGAACTCAGCAATATATTCCGAGTTCATCTCCGCTGCATTTGATTTTACCT
 CAGGATCTATATCCCCTGAACAACAAGTCATCTCTGAGGAAAATGATGCTAAAAAATTAGAGCAAAGTGCT
 CTGAATTGAGAGCAAGCGAAGACTCCGAAGCCATGGATGAAGAATCCAAGGCTCTGAAAGCTGCCGCTGA
 AAAGGCAGATGCCCCGCGCGGTGATTCCCATCCTGGTCTCGCCTGTGACCGCAGCACTGTTCCGGCTG
 GCCTGGACAAGCTGCTGATTATCGGCCCTCGGCTGAGCTCTTCCCATCATCGTTAGCTAGCACTGCGGG
 15 CACGAGGAGACGGCCAGGCCATCGCCTCCTACGGCAGCGCGGTACGCACATCCGGCAGCCCGACCTGAG
 CAGCATTGCGGTGCCGCCGACCCGCAAGTTCCAGGGCTACTACAAGATCGCGCGCCACTACCGCTGGG
 CGCTGGGCCAGGTCTTCCGGCAGTTTCGCTTCCCCGCGGCCGTGGTGGTGGAGGATGACCTGGAGGTGGCC
 CCGGACTTCTTCAGTACTTTCCGGCCACCTATCCGCTGCTGAAGGCCGACCCCTCCCTGTGGTGCCTCTC
 GGCCTGGAATGACAACGGCAAGGAGCAGATGGTGGACGCCAGCAGGCCTGAGCTGCTCTACCGCACCAGCT
 20 TTTTCCCTGGCCTGGGCTGGCTGCTGTTGGCCGAGCTCTGGGCTGAGCTGGAGCCCAAGTGGCCAAAGGCC
 TTCTGGGACGACTGGATGCGGCGGCCGAGCAGCGGCAGGGGCGGGCTGCATACGCCCTGAGATCTCAAG
 AACGATGACCTTTGGCCGCAAGGGTGTGAGCCACGGGCAGTTCTTTGACCAGCACCTCAAGTTTATCAAGC
 TGAACCAGCAGTTTGTGCACTTCACCCAGCTGGACCTGTCTTACCTGCAGCGGGAGGCCTATGACCGAGAT
 TTCCTCGCCCCGCTCTACGGTGCTCCCCAGCTGCAGGTGGAGAAAGTGAGGACCAATGACCGGAAGGAGCT
 25 GGGGGAGGTGCGGGTGAGTATACGGGCAGGGACAGCTTCAAGGCTTTCCGCAAGGCTCTGGGTGTCTATGG
 ATGACCTTAAGTCGGGGGTTCGAGAGCTGGCTACCGGGGTATTGTACCTTCCAGTTCCGGGGCCGCGCT
 GTCCACCTGGCGCCCCCACCAGCTGGGAGGGCTATGATCCTAGCTGGAATTAGCACCTGTGCACTGGAGA
 CCTGCAGGCATGCAAGCTTCGACCATCATCATCATCATCATTGAGTTTGTAGCCTTAGACATGACTGTTCC
 TCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTCTTGCTAGATTCTAATCAAGAGGATGTCAGAATGCCA
 30 TTTGCCTGAGAGATGCAGGCTTCATTTTTGATACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTT
 TGTCATTTTGTTTCTTCTCGTACGAGCTTGCTCCTGATCAGCCTATCTCGCAGCTGATGAATATCTTGTGG
 TAGGGGTTTGGGAAAATCATTGAGTTTGATGTTTTTCTTGGTATTTCCCACTCCTCTTCAGAGTACAGAA
 GATTAAAGTGAGACCTTCGTTTGTGCGGATCCCCACACACCATAGCTTCAAAATGTTTCTACTCCTTTTTT
 ACTCTTCAGATTTTCTCGGACTCCGCGCATCCCGTACCCTTCAAAACACCCAGCACAGCATACTAAA
 35 TTTTCCCTCTTTCTTCTAGGGTGTCGTTAATTACCGTACTAAAGGTTTGGAAAAGAAAAAGAGACC
 GCCTCGTTTTCTTTTTCTTCGTGAAAAAGGCAATAAAAATTTTTATCACGTTTCTTTTTCTT

40 SEQ ID NO: 15
 5' TTCGAAGCTTCGCTAGCTCGGTGTCCCGATGTC

SEQ ID NO: 16
 5' GAATTCGAAGGGAAGATGAGGCTTCGGGAGCC

45 SEQ ID NO: 17
 5' CGTTCGCGACCGGAGGGGCCCCGGCCGCC

SEQ ID NO: 18
 5' TCGATATCAAGCTTAGCTCGGTGTCCCGATGTC

50 SEQ ID NO: 19
 5' GAATTCGAACCTTAAGATGGCCCTCTTTCTCAGTAAG

SEQ ID NO: 20

Human GalT1 cDNA:

ATGAGGCTTCGGGAGCCGCTCCTGAGCGGCGCCGCGATGCCAGGCGCGTCCCTACAGCGGGCCTGCCGCT
GCTCGTGGCCGTCTGCGTCTGGCACCTTGGCGTACCCTCGTTTACTACCTGGCTGGCCGCGACCTGAGCC
GCCTGCCCCAACTGGTCCGAGTCTCCACACCGCTGCAGGGCGGCTCGAACAGTGCCGCCGCCATCGGGCAG
5 TCCTCCGGGGAGCTCCGGACCGGAGGGGCGCGCCGCTCCTCTAGGCGCCTCCTCCAGCCGCGCCC
GGGTGGCGACTCCAGCCCAGTCGTGGATTCTGGCCCTGGCCCCGCTAGCAACTTGACCTCGGTCCCAGTGC
CCCACACCACCGCACTGTCTGCTGCCCGCCTGCCCTGAGGAGTCCCCGCTGCTTGTGGGGCCCCATGCTGATT
GAGTTTAAATGCCTGTGGACCTGGAGCTCGTGGCAAAGCAGAACCCTAAATGTGAAGATGGGCGGCCGCTA
10 TGCCCCCAGGGAAGTGCCTCTCTCCTCACAAGGTGGCCATCATCATTCCATTCCGCAACCGGCAGGAGCACC
TCAAGTACTGGCTATATTATTTGCACCCAGTCCTGCAGCGCCAGCAGCTGGACTATGGCATCTATGTTATC
AACCAGGCGGGAGACACTATATTCAATCGTGCTAAGCTCCTCAATGTTGGCTTTCAAGAAGCCTTGAAGGA
CTATGACTACACCTGCTTTGTGTTTAGTGACGTGGACCTCATTCCAATGAATGACCATAATGCGTACAGGT
GTTTTTTCACAGCCACGGCACATTTCCGTTGCAATGGATAAGTTTGGATTGAGCCTACCTTATGTTTCAAT
15 TTTGGAGGTGTCTCTGCTCTAAGTAAACAACAGTTTCTAACCATCAATGGATTTCCTAATAATTATTGGGG
CTGGGGAGGAGAAGATGATGACATTTTAAACAGATTAGTTTTTAGAGGCATGTCTATATCTCGCCCAAATG
CTGTGGTGGGAGGTGTGCGCATGATCCGCCACTCAAGAGACAAAAAATGAACCCAATCCTCAGAGGTTT
GACCGAATTGCACACACAAAGGAGACAATGCTCTCTGATGGTTTGAACCTCACTCACCTACCAGGTGCTGGA
TGTACAGAGATACCCATTGTATACCCAAATCAAGTGGACATCGGGACACCGAGC

20 SEQ ID NO: 21

Human GalT1 protein:

MRLREPLLGAAMPGLQACRLLVAVCVVHLGVTLVYYLAGRDLRLPQLVGVSTPLQGSNSAAAIGQ
SSGELRTGGARPPPPLGASSQPRPGDSSPVVDSGPGPASNLTSVPVPHTTALSLPACPEESPLLVGPMLI
EFNMPVDLELVAKQNPVVKMGGRYAPRDCVSPHKVAIIIPFRNRQHLKYWLYLHPVLQRQQLDYGIYVI
25 NQAGDTIFNRAKLLNVGFQKALDYDYTCFVFSVDVLIIPMNDHNAYRCFSQPRHISVAMDKFGFSLPYVQY
FGGVSALESKQOFLTNGFPNNYWGWWGGEDDDIFNRLVFRGMSISRPNVAVGRCRMRHSRDKNEPNPQRF
DRIAHTKETMLSDGLNSLTYQVLDVQRYPLYTQITVDIGTPS

30 SEQ ID NO: 22

pBlKanMX4KrehGalT:

CTAGTGCACAAACGAACGTCTCACTTAATCTTCTGTACTCTGAAGAGGAGTGGGAAATACCAAGAAAAACA
TCAAACCTGAATGATTTTCCCAAACCCCTACCACAAGATATTATCATCAGCTGCGAGATAGGCTGATCAGGAG
CAAGCTCGTACGAGAAGAAACAAAATGACAAAAAAATCCTATACTATATAGGTTACAAATAAAAAAGTAT
35 CAAAAATGAAGCCTGCATCTCTCAGGCAAATGGCATTCTGACATCCTCTTGATTAGAATCTAGCAAGACCG
GTCTTCTCGTAAGTGCCCAACTTGAACTGAGGAACAGTCATGTCTAAGGCGAATTATCAAGCTTAGCTCGG
TGTCCCGATGTCCACTGTGATTTGGGTATACAATGGGTATCTCTGTACATCCAGCACCTGGTAGGTGAGTG
AGTTCAAACCATCAGAGAGCATTGTCTCTTTGTGTGTGCAATTGGGTCAAACCTCTGAGGATTGGGTTCA
40 TTTTTTTTGTCTCTTGAGTGGCGGATCATGCGACACCTCCCGACCACAGCATTGCGGCGAGATATAGACAT
GCCTCTAAAACTAATCTGTTAAAAATGTATCATCTTCTCCTCCCGAGCCCCAATAATTATTAGGAAATC
CATTGATGGTTAGAACTGTTGTTTACTTAGAGCAGAGACACCTCCAAAATACTGAACATAAGGTAGGCTG
AATCCAAACTTATCCATTGCAACGGAAATGTGCCGTGGCTGTGAAAAACACCTGTACGCATTATGGTCATT
CATTGGAATGAGGTCCACGTCACTAAACACAAAGCAGGTGTAGTCATAGTCCTTCAAGGCTTCTTGAAAGC
CAACATTGAGGAGCTTAGCACGATTGAATATAGTGTCTCCCGCCTGGTTGATAACATAGATGCCATAGTCC
45 AGCTGCTGGCGCTGCAGGACTGGGTGCAAATAATATAGCCAGTACTTGAGGTGCTCCTGCCGGTTGCGGAA
TGGAATGATGATGGCCACCTTGTGAGGAGAGACGCGAGTCCCTGGGGGCATAGCGGCCGCCCATCTTACAT
TTGGGTTCTGCTTTGCCACGAGCTCCAGGTCCACAGGCATGTTAAACTCAATCAGCATGGGGCCCAAGC
AGCGGGGACTCCTCAGGGCAGGCGGGCAGCGACAGTGCGGTGGTGTGGGGCACTGGGACCGAGGTCAAGTT
GCTAGCGGGGCCAGGGCCAGAATCCACGACTGGGCTGGAGTCGCCACCCGGGCGCGGCTGGGAGGAGGCGC
50 CTAGAGGAGGCGGCGGCGGGCCCCCTCCGGTCGCCGGCGGGGCATCTGCCTTTTTCAGCGGCAGCTTTTCA
GCCTTGGATTCTTCATCCATGGCTTCGGAGTCTTCGCTTGCCTCTGAATTCAGAGCACTTTGCTCTAATTT
TTTAGCATCATTTTCTCAGAGATGACTTGTGTTTCAAGGGATATAGATCCTGAGGTAAATCAAATGCAG
CGGAGATGGAACCTCGAATATATTGCTGAGTCTACTGTTTGAATTCAATGTTAGGAGGAGAACAAATAACC
GCACCTGCAATGACGGTAAATCTCAACAGTCTCTTACTGAGAAAGAGGGCCATCTTAAGTTTCAATAATTA

GTTGTTTTTTGTATCTTCTCAAGTTGTGCGTTAAAAGTCGTTAAAATCAAAAGCTTGTCAATTGGAACCAGTC
 GCAATTATGAAAGTAAGCTAATAATGATGATAAAAAAAGGTTTAAAGACAGGGCAGCTTCCTTCTGTTTA
 TATATTGCTGTCAAGTAGGGGTTAGAACAGTTAAATTTTGTATCATGAACGTTAGGCTATCAGCAGTATTCC
 CACCAGAATCTTGGAAGCATACAATGTGGAGACAATGCATAATCATCCAAAAGCGGGTGTTCCTTCCCATTT
 5 GCGTTTTCGGCACAGGTGCACCGGGGTTGAGAAGCGATAGAGAGACTGCGCTAAGCATTAAATGAGATTATTT
 TTGAGCATTTCGTCAATCAATACCAACAAGACAAACGGTATGCCGACTTTTGGAAAGTTTCTTTTTGACCAA
 CTGGCCGTTAGCATTTCAACGAACCAAACTTAGTTCATCTTGGATGAGATCACGCTTTTGTTCATATTAGGT
 TCCAAGACAGCGTTTAAACTGTGAGTTTGGGCCATTTGGGGAACATGAACTATTTGACCCACACTCAG
 AAAGCCCTCATCTGGAGTGATGTTGCGGTGTAATGCCGAGCTTGTTCGATTTCGGAAATAAAACAAACATGAA
 10 CCTCGCCAGGGGGGCCAGGATAGACAGGCTAATAAAGTCATGGTGTAGTAGCCTAATAGAAGGAATTGGA
 ATGAGCGAGCTCCAATCAAGCCCAATAACTGGGCTGGTTTTTCGATGGCAAAAGTGGGTGTTGAGGAGAAG
 AGGAGTGGAGGTCTGCGTTTGCAACGGTCTGCTGCTAGTGTATCCCCTCTGTTGCGTTTGGCATTATG
 TGTGAGAAATGGACCTGTGGATGTGCGATGGCAAAAAGGTTTCATTCAACCTTTGCTTTGGATGTTAGCT
 AGCCGGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGTATTGGGCGCTCTTCGCTTC
 15 CTCGCTCACTGACTCGCTGCGCTCGGTGCTTCGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAA
 TACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGG
 AACCGTAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCG
 ACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCC
 TCGTGCGCTCTCCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTG
 20 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGT
 GCACGAACCCCCGTTGAGCCCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAA
 GACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCT
 ACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCT
 GAAGCCAGTTACCTTCGGAAGAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCACCGCTGGTAGCGGTG
 25 GTTTTTTTGTTTGAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCT
 ACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGAT
 CTTACCTAGATCCTTTTAAATTAATAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGT
 CTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTT
 GCCTGACTCCCCGTGCTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGAT
 30 ACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCA
 AAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGT
 TCGCCAGTTAATAGTTTGCAGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGCTTTGG
 TATGGCTTCATTACGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAG
 CGGTTAGCTCCTTCGGTCTCCGATCGTTTGTGAGAAGTAAGTTGGCCGAGTGTTATCACTCATGGTTATG
 35 GCAGCACTGCATAATTCTCTTACTGTGATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAAC
 CAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCG
 CGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATC
 TTACCGCTGTTGAGATCCAGTTCGATGTAACCACTCGTGACCCAACTGATCTTCAGCATCTTTTACTTT
 CACCAGCGTTTCTGGGTGAGCAAAAAAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA
 40 AATGTTGAATACTCATACTCTTCTTTTCAATGCTTCTAGTGGTAGGAATTAATTCTGTACCGTTTACA
 GAAGGACGACTCTTGATGCGCCAACACAGTGACAATAGACATAGAGGAAATGACAAAAGGATTATGCGAG
 GATGCTGCTGGAGACGATTCAAAGTTTAGTTTAGAAAGGTCTCCATTTATGCTGATAGAATACTAGATAC
 CCGTGAACTTTGTCTCAGGAGATCCGCATCAGACGAAGGATGTTCCGACCTGCAATAATCGAAGAAGAGA
 CCCCTAGGCAGTTGGTGAGCTTACATGAGAAGTCTAAACTATCTTGGACCCGCTGGTTTTATAAAGGGTTC
 45 GTTAGGAATGCGTTAACTACCATTCAGCAACATCCGTGGGGCTTCTGGTGTGTTGAAATACTGCGTCAAAA
 ATTGAGCGATGAAATTGAAGATCGATTGAGTTGAATCGCCGAAACAATTGATCCCCTGTACATACTTGTA
 ATTTACCTCAGAATGGGTAAATTAAGGCGCGCCAGATCTGTTTAGCTTGCCCTCGTCCCCGCCGGGTCAACC
 GGCCAGCGACATGGAGGCCCAGAATACCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGGCATGATGTG
 ACTGTCGCCCCGTACATTTAGCCCATACATCCCATGTATAATCATTGTCATCCATACATTTTGTGCGCCG
 50 ACGGCGCAAGCAAAAATTACGGCTCTGCTGCGAGACCTGCGAGCAGGAAACGCTCCCCTCACAGACGC
 GTTGAATTGTCCCCACGCGCGCCCTGTAGAGAAATATAAAGGTTAGGATTTGCCACTAGAGTTCTTCT
 TTCATATACTTCTTTTAAAATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAAACCACTGG
 GTAAGGAAAAGACTCACGTTTTCAGAGCCGCGATTAAATCCAACATGGATGCTGATTTATATGGGTATAAA
 TGGGCTCGCGATAATGTGCGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGA

5 GTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGC
 TGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACC
 ACTGCGATCCCCGGCAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTAGGTGAAAATATTGTTGA
 TCGCTGGCAGTGTTCTGCGCCGGTTGCATTGATTCTGTTTGTATTGTCTTTTAAACAGCGATCGCG
 10 TATTTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAG
 CGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTGAGT
 CGTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATG
 TTGGACGAGTCGGAATCGCAGACCGATAACCAGGATCTTGCCATCCTATGGAAGTGCCTCGGTGAGTTTTCT
 CCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCA
 15 TTTGATGCTCGATGAGTTTTCTAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACTTGTCATT
 TGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAAATGTTAGCGTGATTTATATTTTTTTTCGCC
 TCGACATCATCTGCCCAGATGCGAAGTTAAGTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAATG
 CTGGTCGCTATACTGCTGTCGATTGATACTAACGCCGCCATCCAGTGTCGAAAACGAGCTCGAATTCATC
 GATGATATCAGATCCA